a. $-\sum_{w\in Vocab} y_w \log(\hat{y}_w) = -(y_1\log(\hat{y}_1) + \cdots + y_o \log(\hat{y}_o) + \cdots + y_{|V|}\log(\hat{y}_{|V|}))$

$\because y_w = \begin{cases} 1 & w=o \\ 0 & w\neq o \end{cases}$

$\therefore$ 原式 $= -y_o \log(\hat{y}_o) = -\log(\hat{y}_o)$

b. $\dfrac{\partial J_{naive\text{-}softmax}(v_c, o, U)}{\partial v_c} = \dfrac{\partial}{\partial v_c}\left(-\log \dfrac{\exp(u_o^T \cdot v_c)}{\sum_{w\in Vocab}\exp(u_w^T \cdot v_c)}\right)$

原式 $= \dfrac{\partial}{\partial v_c}\left(-\log \exp(u_o^T v_c) + \log \sum_{w\in Vocab} \exp(u_w^T \cdot v_c)\right)$

$= \dfrac{\partial}{\partial v_c}\left(-u_o^T v_c) + \log \sum_{w\in Vocab} \exp(u_w^T v_c)\right)$

$= -u_o + \dfrac{\frac{\partial}{\partial v_c}\left(\sum_{w\in Vocab}\exp(u_w^T \cdot v_c)\right)}{\cancel{\log}\cdot \sum_{w\in Vocab}\exp(u_w^T v_c)}$

$= -u_o + \dfrac{\sum_{w\in Vocab}(\exp(u_w^T v_c)\cdot u_w)}{\sum_{w\in Vocab}\exp(u_w^T v_c)}$

$= -u_o + \sum_{w\in Vocab} \hat{y}_w u_w$

$= \sum_{w\in Vocab}(\hat{y}_w u_w - y_w u_w)$

$= \sum_{w\in Vocab}(\hat{y}_w - y_w) u_w = U(\hat{y} - y)$

c. $\dfrac{\partial J_{naive\text{-}softmax}(v_c, o, U)}{\partial u_w} = \dfrac{\partial\left(-\log\exp(u_o^T v_c) + \log \sum_{w\in Vocab}\exp(u_w^T v_c)\right)}{\partial u_w}$

$= -\dfrac{\partial u_o^T}{\partial u_w}\cdot v_c + \dfrac{\sum_{w\in Vocab}(\exp(u_w^T \cdot v_c) v_c)}{\sum_{w\in Vocab}\exp(u_w^T v_c)}$

$\dfrac{\partial u_o^T}{\partial u_w} = \begin{cases} 1 & w=o \\ 0 & w\neq o \end{cases} \quad \therefore \dfrac{\partial u_o^T}{\partial u_w} = y_w = \sum_{w\in Vocab} y_w$

$\therefore$ 原式 $= \sum_{w\in Vocab}\left[-y_w v_c + \dfrac{\exp(u_w^T v_c)}{\sum_{w\in Vocab}\exp(u_w^T \cdot v_c)}\cdot v_c\right)$

$= \sum_{w\in Vocab} v_c(\hat{y}_w - y_w)$

$= \begin{cases} v_c(\hat{y}_w - 1) & w=o \\ v_c \hat{y}_w & w\neq o \end{cases}$

d. $\frac{\partial J_{naive\text{-}softmax}(v_c, o, U)}{\partial U} = \left[\frac{\partial J}{\partial u_1}; \frac{\partial J}{\partial u_2} \cdots \frac{\partial J}{\partial u_{|V|}} \right]$

e. $$\frac{\partial \sigma}{\partial x} = \frac{e^x(e^x+1) - e^x e^x}{(e^x+1)^2} = \frac{e^x}{(e^x+1)^2} = \sigma(x)\cdot(1-\sigma(x))$$

f.

(1) $$\frac{\partial J_{neg\text{-}sample}(v_c, o, U)}{\partial v_c} = \frac{\partial}{\partial v_c}\left[-\log(\sigma(u_o^T v_c)) - \sum_{k=1}^{K} \log(\sigma(-u_k^T v_c))\right]$$

$$= \left(-\frac{1}{\sigma(u_o^T v_c)}\right)\left(u_o \sigma(u_o^T v_c)(1-\sigma(u_o^T v_c))\right) - \sum_{k=1}^{K} \frac{-u_k \sigma(-u_k^T v_c)(1-\sigma(-u_k^T v_c))}{\sigma(-u_k^T v_c)}$$

$$= -u_o(1-\sigma(u_o^T v_c)) + \sum_{k=1}^{K} u_k (1-\sigma(-u_k^T v_c))$$

$$= u_o(\sigma(u_o^T v_c)-1) - \sum_{k=1}^{K} u_k(\sigma(-u_k^T v_c)-1)$$

(2) $$\frac{\partial J_{neg\text{-}sample}(v_c, o, U)}{\partial u_o} = \frac{\partial}{\partial u_o}\left[-\log(\sigma(u_o^T v_c)) - \sum_{k=1}^{K} \log(\sigma(-u_k^T v_c))\right]$$

$$= \frac{\partial}{\partial u_o}\left[-\log(\sigma(u_o^T v_c))\right] - \frac{\partial}{\partial u_o}\left[\sum_{k=1}^{K} \log(\sigma(-u_k^T v_c))\right]$$

$$= \frac{\partial}{\partial u_o}\left[-\log(\sigma(u_o^T v_c))\right]$$

$$= -\frac{1}{\sigma(u_o^T v_c)} \left(\sigma(u_o^T v_c)(1-\sigma(u_o^T v_c)) v_c\right)$$

$$= v_c(\sigma(u_o^T v_c) - 1)$$

(3) $$\frac{\partial J_{neg\text{-}sample}(v_c, o, U)}{\partial u_k} = 0 - \frac{\partial}{\partial u_k}\left[\sum_{k=1}^{K} \log(\sigma(-u_k^T v_c))\right]$$

$$= \frac{-v_c \sigma(-u_k^T v_c)(1-\sigma(-u_k^T v_c))}{\sigma(-u_k^T v_c)} = v_c(1-\sigma(-u_k^T v_c))$$

g.

$$\frac{\partial J_{neg\text{-}sample}(v_c; o, U)}{\partial u_k} = \frac{\partial}{\partial u_k}\left[-\log(\sigma(u_o^T v_c)) - \sum_{k=1}^{K}\log(\sigma(-u_k^T v_c))\right]$$

$$= \frac{\partial}{\partial u_k}\left[-\log(\sigma(u_o^T v_c))\right] - \frac{\partial}{\partial u_k}\left[\sum_{k=1}^{K}\log(\sigma(-u_k^T v_c))\right]$$

$$= 0 - \frac{\partial}{\partial u_k}\left[\sum_{k=1}^{K}\log(\sigma(-u_k^T v_c))\right]$$

$$= -\left[\sum_{i\in\{1,\cdots,K\}: w_i = w_k}\frac{\partial}{\partial u_k}\log(\sigma(-u_i^T v_c)) + \sum_{i\in\{1,\cdots,K\}: w_i \neq w_k}\frac{\partial}{\partial u_k}\log(\sigma(-u_{i, w_i\neq w_k}^T v_c))\right]$$

$$= -\sum_{i\in\{1,\cdots,K\}: w_i = w_k}\frac{\partial}{\partial u_k}\log(\sigma(-u_i^T v_c))$$

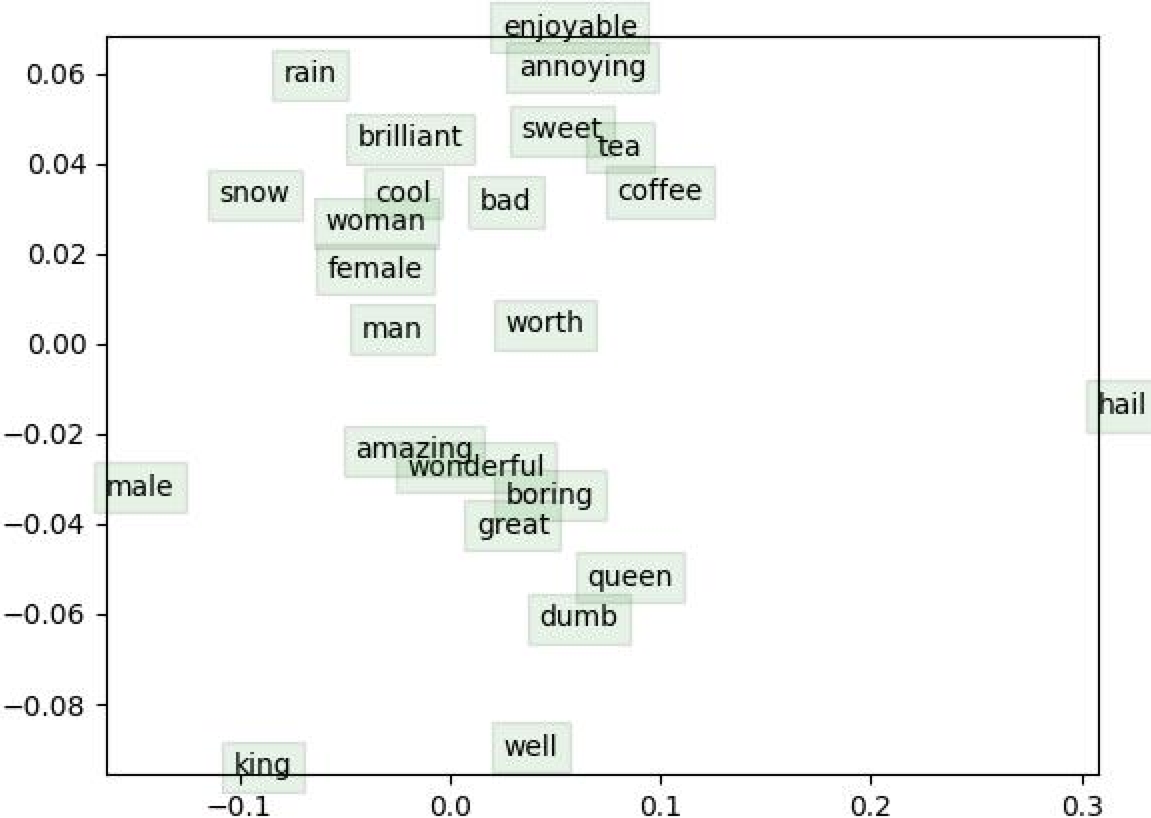
$$= \sum_{i\in\{1,\cdots,K\}: w_i = w_k} v_c(1-\sigma(-u_k^T v_c))$$

h.

(1) $$\frac{\partial J_{skip\text{-}gram}(v_c, w_{t-m}, \cdots w_{t+m}, U)}{\partial U} = \sum_{\substack{-m\le j\le m \\ j\neq 0}} \frac{J(v_c, w_{t+j}, U)}{\partial U}$$

(2) $$\frac{\partial J_{skip\text{-}gram}(v_c, w_{t-m}, \cdots, w_{t+m}, U)}{\partial v_c} = \sum_{\substack{-m\le j\le m \\ j\neq 0}} \frac{J(v_c, w_{t+j}, U)}{\partial v_c}$$

(3) $$\frac{\partial J_{skip\text{-}gram}(v_c, w_{t-m}, \cdots, w_{t+m}, U)}{\partial v_{w}\ w\neq c} = 0$$



enjoyable
annoying
rain
brilliant
sweet
tea
snow
cool
bad
coffee
woman
female
man
worth
hail
amazing
wonderful
male
boring
great
queen
dumb
well
king
0.06
0.04
0.02
0.00
−0.02
−0.04
−0.06
−0.08
−0.1
0.0
0.1
0.2
0.3